

**Debate Dialog for  
News Question Answering System ‘NetTv’  
~Debate Based on Claim and Reason Estimation~**

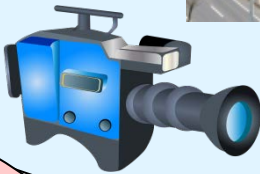
Rikito Marumoto, Katsuyuki Tanaka,  
Tetsuya Takiguchi, Yasuo Ariki

# Background

We have questions in everyday life.

# Real world

We need an assistant to understand them deeper and wider.

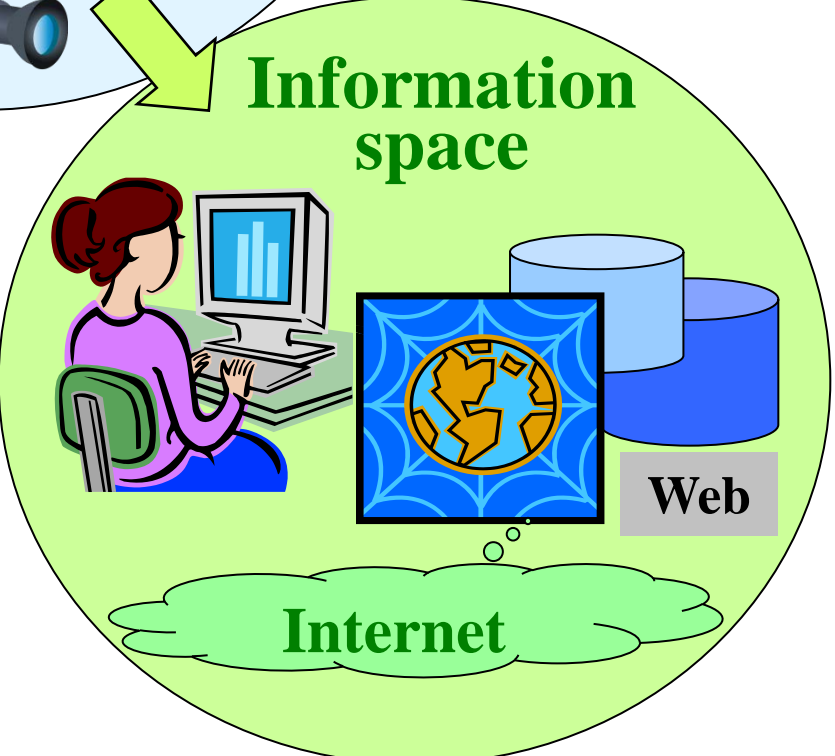
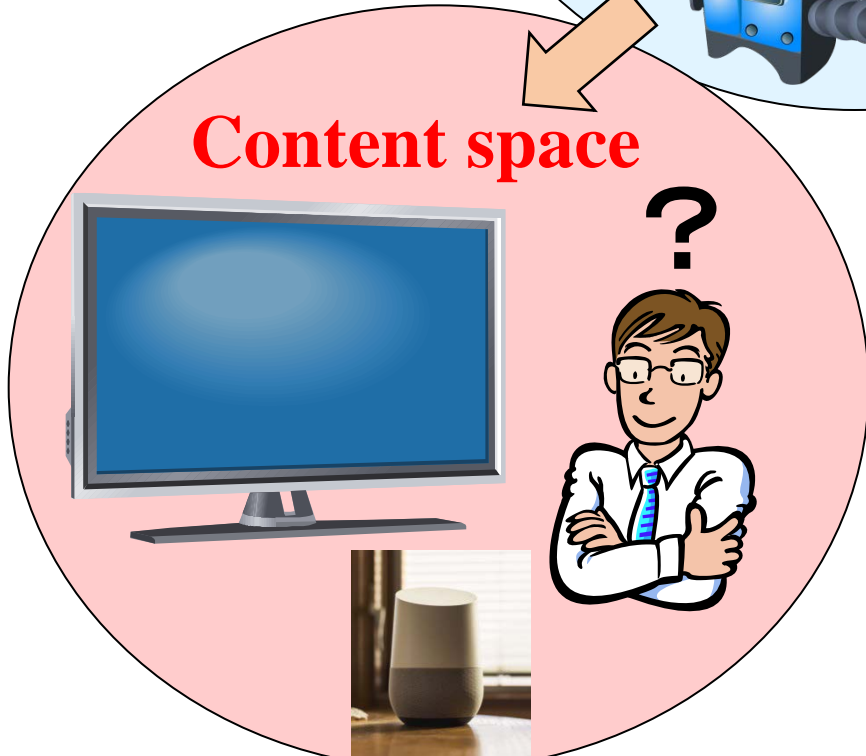


What is Casino bill?

What causes Tsunami?

# Content space

# Information space

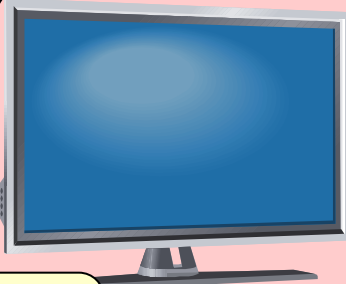


**Research purpose**

Show me news on xx.  
What is Casino bill?

For deeper and wider understanding.

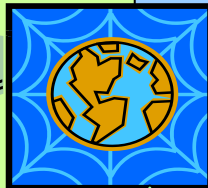
**Content space**



About  
xxx...

**Net TV**

**Information space**



We

**Internet**

**Speech recognition**

Dictionary update  
Grammar

**QA module**

**Q classification**

- Factoid Q
- Definition Q
- Why Q
- How Q

**Answer retrieval**

**Debate module**

**News indexing**

- Web spider
- Html parser
- Morphological analysis

**NetNews module**

**Net news database**

- Newspaper\_URL
- Movie\_URL
- Time, Headline

# Debate system

## [1] Classifying Stance and Reasons, Predicting Winners

### Classifying Positions: Support or Oppose?

Akiko.Murakami, Rudy.Raymond, “Support or Oppose? Classifying Positions in Online Debates from Reply Activities and Opinion Expressions,” in Proceedings of the 23rd ICCL, pp.869-875, 2010.

### Identifying and Classifying Reasons

Kazi Saidul Hasan and Vincent Ng , “Why are You Taking this Stance? Identifying and Classifying Reasons in Ideological Debates,” Proceedings of the EMNLP, pages 751–762, October 25-29, 2014.

### Predicting Debate Winners

Peter Potash and Anna Rumshisky, “Towards Debate Automation: a Recurrent Model for Predicting Debate Winners,” Proceedings of the EMNLP, pages 2465–2475 Copenhagen, Denmark, September 7–11, 2017.

## [2] Argument Visualization and Generation

### Argument Generation System

Users can specify a topic and a stance (agree/disagree).

The system outputs three argument paragraphs based on “values”.

Each value is regarded as a viewpoint of the generated argument.

### OpinionReader: Argument Visualization

System summarizes and visualizes the arguments for a target topic in two-dimension of polarity and importance.

# Argument Generation System

Users can specify a **topic** and a **stance** (agree/disagree).  
The system outputs three **argument** paragraphs based on “**values**”.  
The “value” represents what is important in human’s life, what is harmful in communities. The value dictionary formulates a set of values.  
Each value is regarded as a **viewpoint** of the generated argument.

Misa Sato, Kohsuke Yanai, Toshihiko Yanase, Toshinori Miyoshi, Makoto Iwayama, Qinghua Sun, Yoshiki Niwa, “End-to-end Argument Generation System in Debating,” Proceedings of ACL-IJCNLP 2015 System Demonstrations, pages 109–114, Beijing, China, July 26-31, 2015.

**Stance  
(agree)**

The screenshot displays the interface of the Argument Generation System. At the top left, there is a 'Start' button and a 'Stance' selector with 'Agree' selected. The 'Topic' input field contains the text 'This house should ban smoking in public spaces'. Below the topic, there is a timer showing '180 sec.'. To the right, a statistics panel lists: documents: 3500, sentences: 944, assertion: 433, example: 496, and causal effect: 10. Further right is a bar chart titled '# of sentences' with categories: disease, pollution, poverty, health, cost, and environment. The 'disease' category has the highest count, around 350. Below these elements, a large text box shows a generated argument: 'It increases poverty; On a country basis, smoking increases poverty levels by raising health costs, crippling members of the work-force and harming the environment.' To the right of this text are four progress bars for 'motion analysis', 'value selection', 'retrieval', and 'ordering', all showing high completion. At the bottom, three columns represent different values: 'poverty', 'pollution', and 'disease'. Each column contains a paragraph of text generated from that value.

Value	Generated Argument Paragraph
poverty	It increases poverty; On a country basis, smoking increases poverty levels by raising health costs, crippling members of the work-force and harming the environment. The report called for a worldwide response to the health problems caused by household smoke, and argued that the issue should be formally recognized by the UN in
pollution	It caused pollution; Smoke from oilwell fires and burning oil trenches had caused local air pollution and soil contamination. An expert said a smoking ban would force customers to file outdoors every 10 minutes to light up, a situation that could create noise pollution in residential areas. The Jakarta city council approved the smoking ban as part of a law
disease	It causes disease; Nicotine is addictive. The tobacco industry has had to accept government and legal advice that tobacco can cause disease but has championed smoking as a matter of personal choice. Surveys show that smoking can not only cause the similar hazard on women as on men, but also lead to

**Values &  
#sentences**

**Three  
argument  
paragraphs**

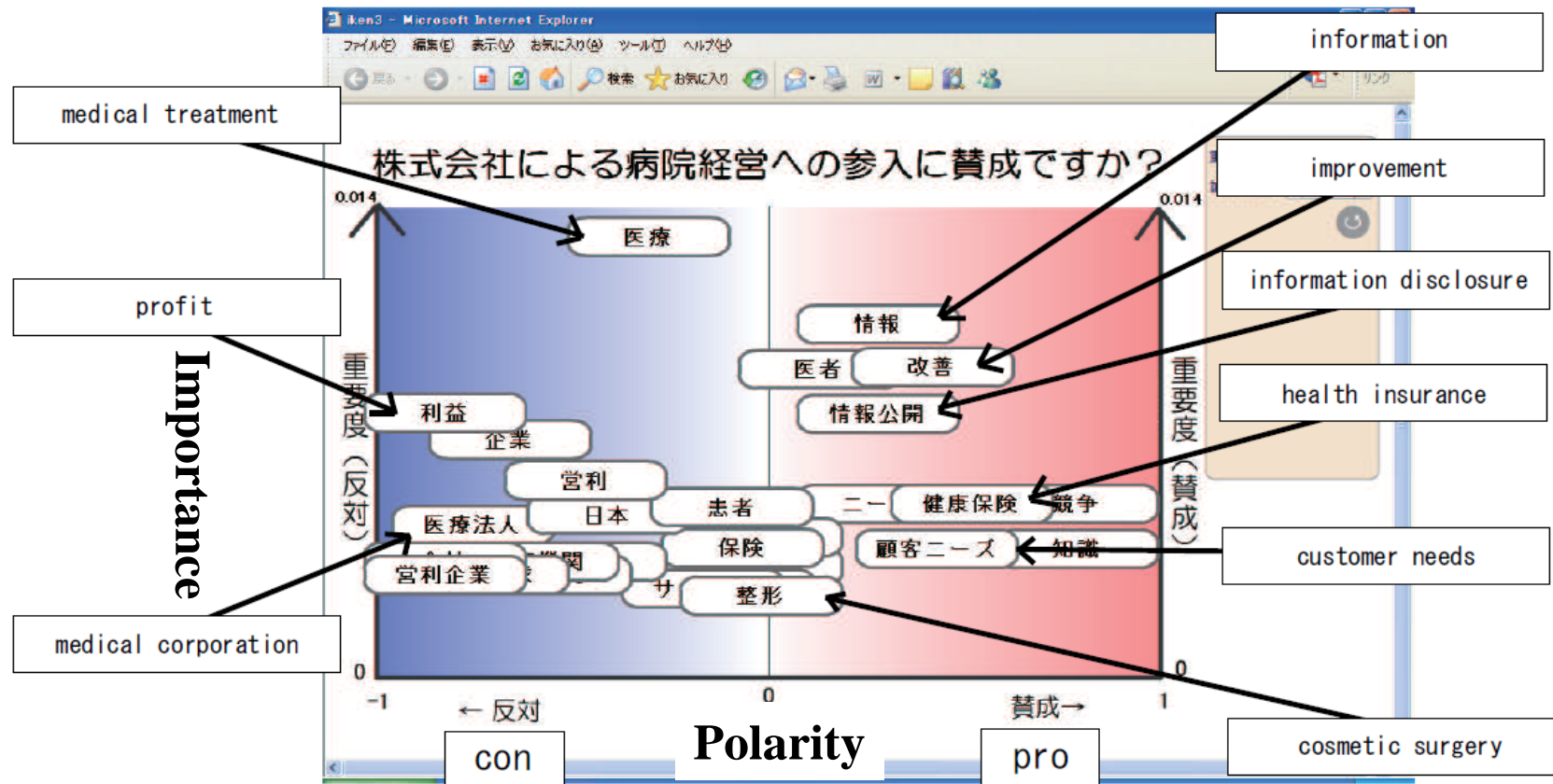
# OpinionReader : Argument Visualization

System summarizes and visualizes the **arguments** for a target **topic** in two-dimension of **polarity** and **importance**.

Users can catch the arguments, choosing a more reasonable **viewpoint** for **decision making**.

A. Fujii and T. Ishikawa, "A system for summarizing and visualizing arguments in subjective documents: Toward supporting decision making," Proc. COLING-ACL Workshop on Sentiment and Subjectivity in Text, pp.15-22, 2006.

## Topic: Privatization of hospitals by joint-stock companies?



# Debate components

*Topics*

This house should ban smoking in public spaces.

*Viewpoints (values)*

[stance(polarity)/importance]

Poverty, pollution, disease, health, cost

*Evidences(fact,logic)*

[pos/neg  $\Rightarrow$  values up/down]



*Reasons*

Smoking causes cancers  $\Rightarrow$  values down

*Arguments*

[topic, stance, reasons]



*Debate*

*Decision making*

[stance-decision, for the topic, based on reasons]

## Argument Generation System

Users can specify a **topic** and a **stance** (agree/disagree).

The system outputs three **argument** paragraphs based on “**values**”.

The “value” represents what is important in human’s life, what is harmful in communities. The value dictionary formulates a set of values.

Each value is regarded as a **viewpoint** of the generated argument.

## OpinionReader: Argument Visualization

System summarizes and visualizes the **arguments** for a target **topic** in two-dimension of **polarity** and **importance**.

Users can catch the arguments, choosing a more reasonable **viewpoint** for **decision making**.



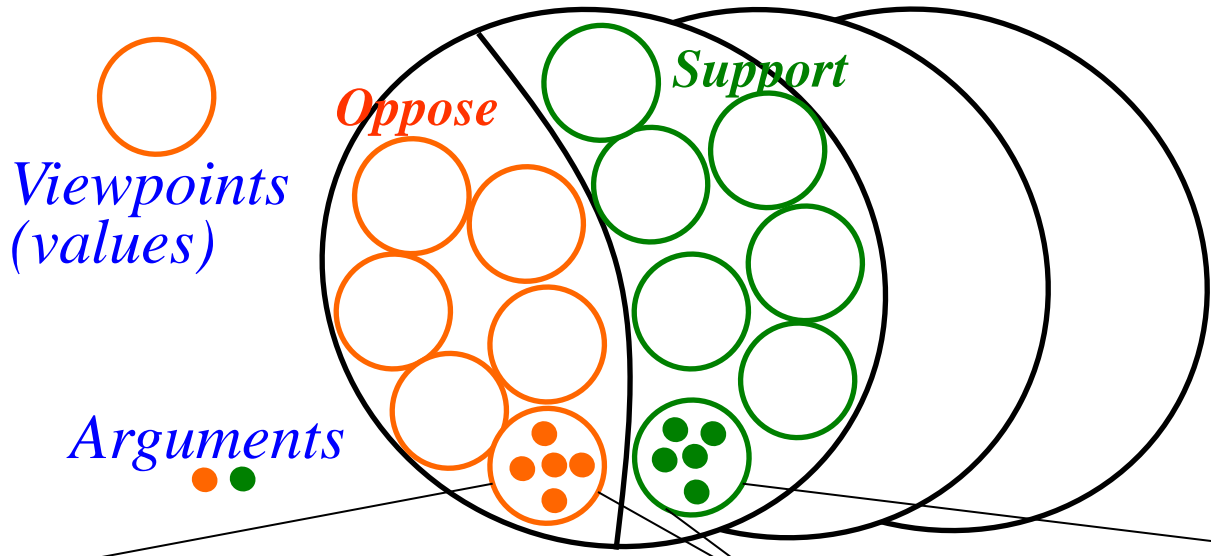
# Debate components

- Topics* This house should ban smoking in public spaces.
- Viewpoints (values)* [stance(polarity)/importance]  
Poverty, pollution, disease, health, cost
- Evidences(fact,logic)* [pos/neg  $\Rightarrow$  values up/down]  $\rightarrow$  *Reasons*  
Smoking causes cancers  $\Rightarrow$  values down
- Arguments* [topic, stance, reasons]  $\rightarrow$  *Debate*

## Debate database

### Topics

### Arguments & viewpoints



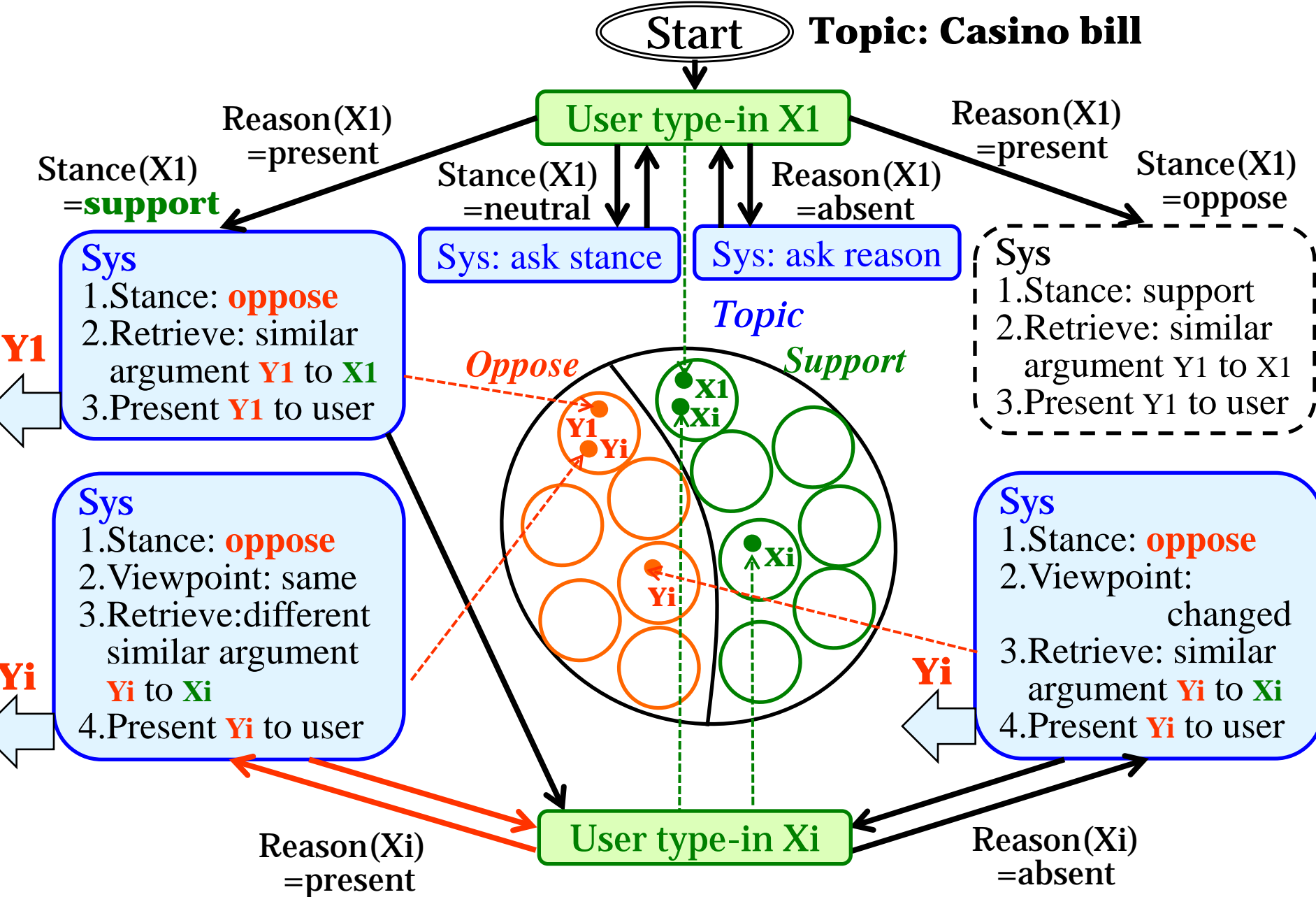
- Documents search for each topic
- Arguments extraction from the documents
- Filtering arguments with high score
- Viewpoint clustering by LDA

ID	Text	Class	Vector
0	Disagree ...	2	[0,1,1,...,0,1]
1		2	

ID	Text	Class	Vector
0	Disagree ...	5	[0,0,0...1,1]
1		5	



# Debate State Diagram



# Stance & Reason estimation

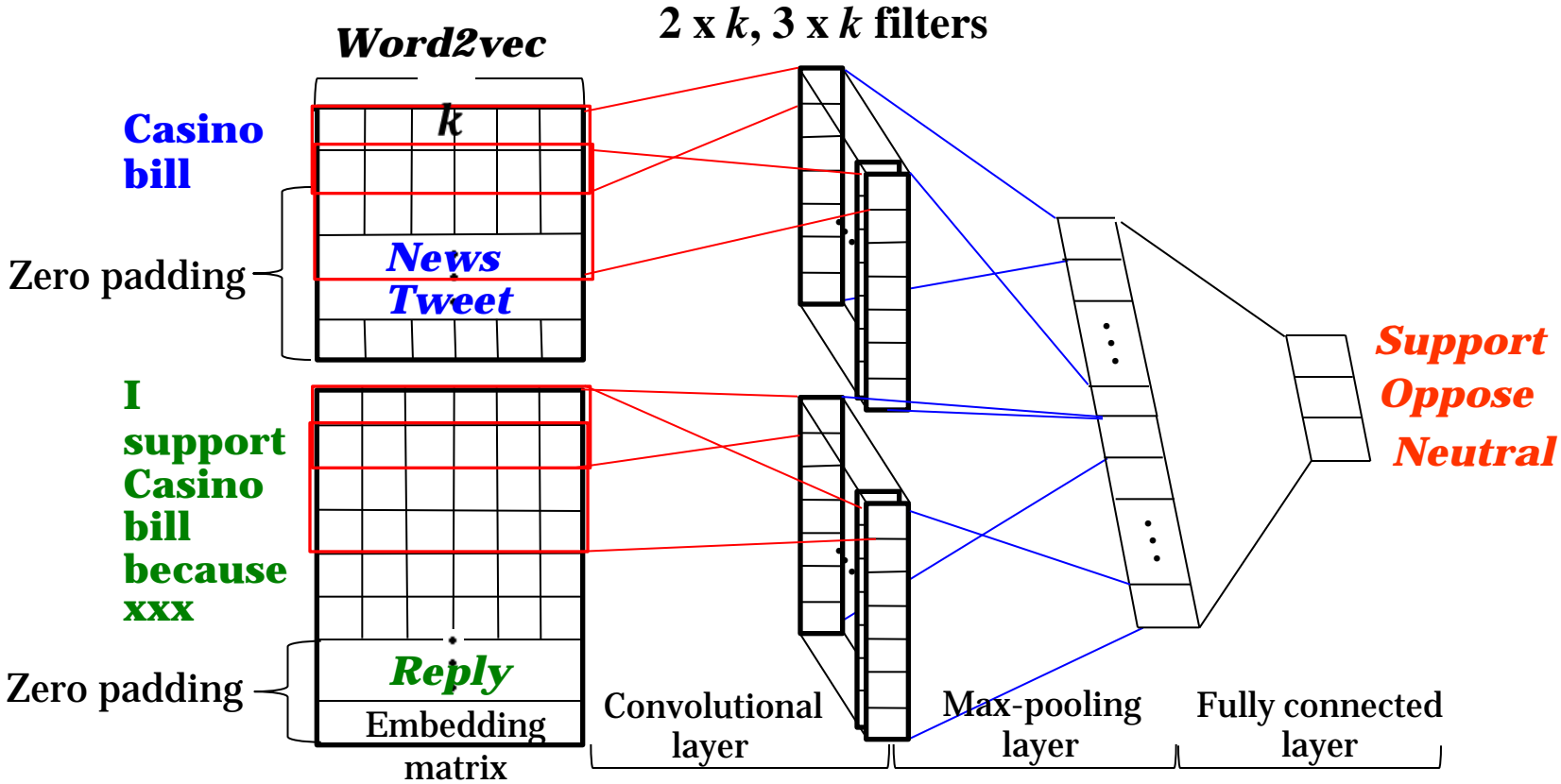
*Training data:* news Tweets and the paired user Replies.

*Input:* word2vec word embedding learned with Wikipedia.

*Output:* Stance **support/oppose/neutral** Reason **present/absent**

*CNN model:*

Hongjie Shi, Takashi Ushio, Mitsuru Endo, katsuyoshi Yamagami and Noriaki Horii, "A Multichannel Convolutional Neural Network For Cross-language Dialog State Tracking," *IEEE Workshop on Spoken Language Technology*, 2016.



# Stance & Reason estimation

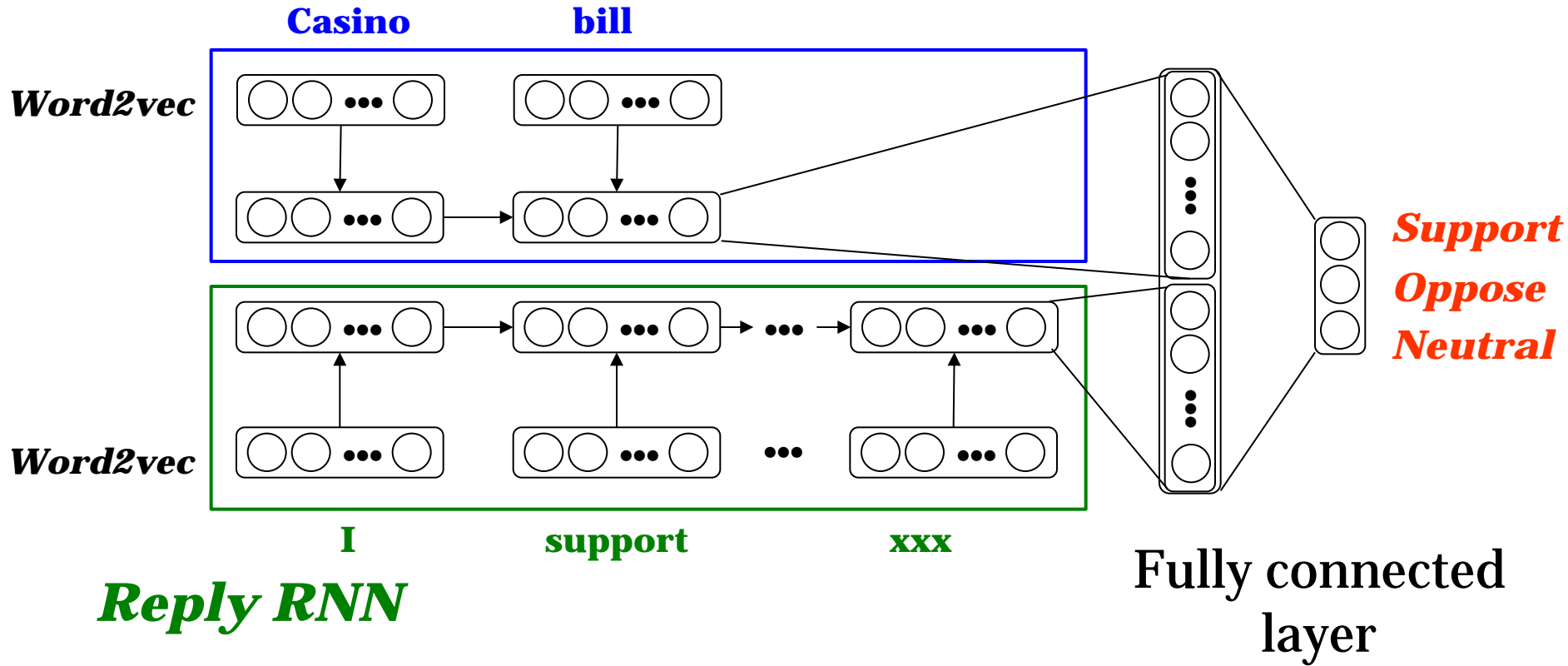
*Training data:* news Tweets and the paired user Replies.

*Input:* word2vec word embedding learned with Wikipedia.

*Output:* Stance **support/oppose/neutral** Reason **present/absent**

*RNN model:*

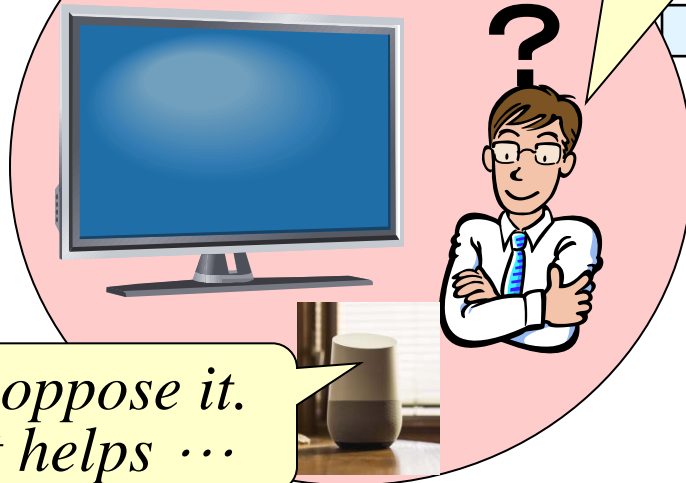
## News Tweet RNN



# Net TV

*I support casino bill because it xxxxx.*

## Content space



*I oppose it. It helps ...*

# Debate system for deeper and wider understanding

Text input

Text output

Language understanding

Debate management

## Word extraction

- Morphological analysis

## Stance estimation

- Support
- Oppose
- Neutral

## Reason estimation

- Present
- Absent

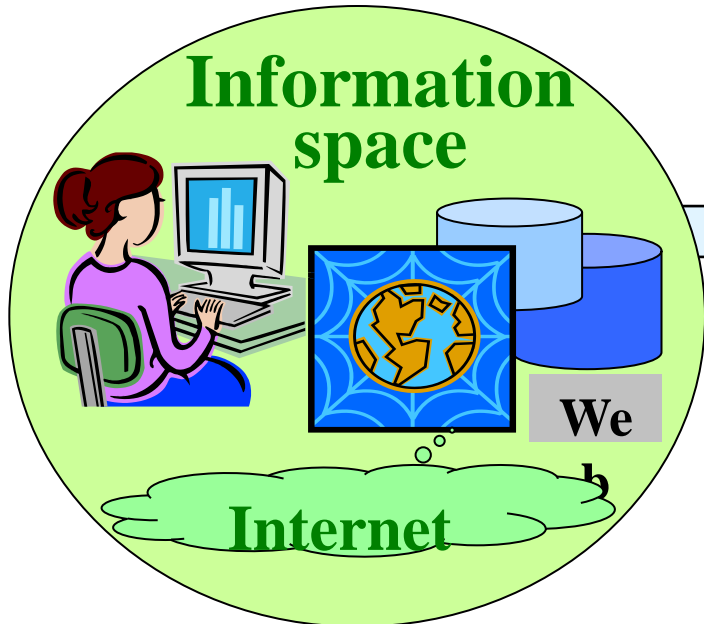
## Argument retrieval

- Opposite stance
- Similar viewpoint

## Argument & view

- Initial search
- Re-search
- Filtering
- LDA

## Information space



Internet

ID	Text	Class	Vector	ID	Text	Class	Vector
0	Disagree ...	2	[0,1,1,...,0,1]	0	Disagree ...	5	[0,0,0... ,1,1]
1		2		1		5	

# Experiment

## *Experimental data:*

News Tweets(topic) of the Sankei newspaper official account (@Sankei\_news) and the paired user Replies acquired by Twitter API.

## *Stance estimation (support/oppose/neutral)*

*Training data:* 1486 sentences (support: 458, oppose: 831, neutral: 831)

*Test data :* 150 sentences (support: 50, oppose: 50, neutral: 50)

## *Reason estimation (Present/Absent)*

*Training data :* 1301 sentences (present: 218, absent: 1083)

*Test data :* 100 sentences (present: 50, absent: 50)

## *Experimental condition:*

### *RNN hyper parameters*

	<b>Stance</b>	<b>Reason</b>
Dropout rate	0.5	0.4
# of hidden nodes	-	64
Learning rate	0.0001	0.001

### *CNN hyper parameters*

	<b>Stance</b>	<b>Reason</b>
Dropout rate	0.5	0.4
# of filters (1 x k)	-	64
# of filters (2 x k)	64	64
# of filters (3 x k)	64	-
Learning rate	0.0001	0.001
Weight decay	0.0001	0.0001

# Experiment

## *Experimental result:*

*Accuracy(%)*

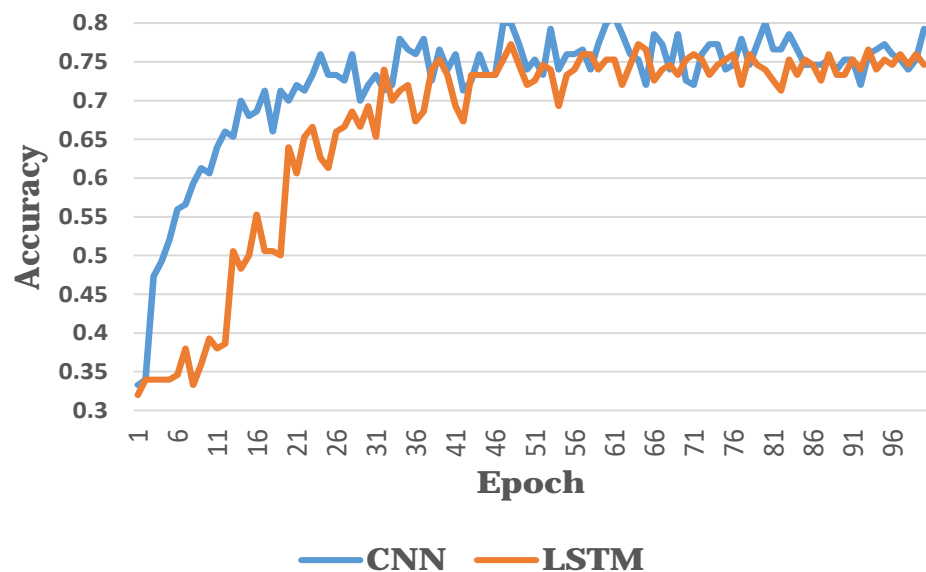
	<b>Stance</b>	<b>Reason</b>
CNN	80.0	80.0
RNN	78.0	80.0

*Training time for 1 epoch(sec)*

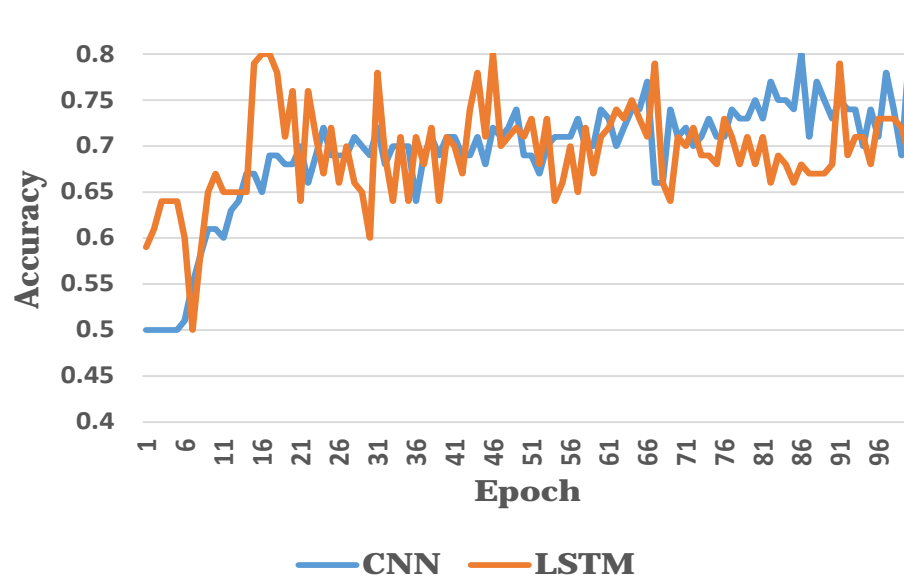
	<b>Stance</b>	<b>Reason</b>
CNN	7.41	6.808
RNN	30.32	30.02

*Accuracy as a function of epoch*

## **Stance**



## **Reason**



# Experiment

# Subjective evaluation

## Experimental condition:

Number of arguments  
Collected in debate database

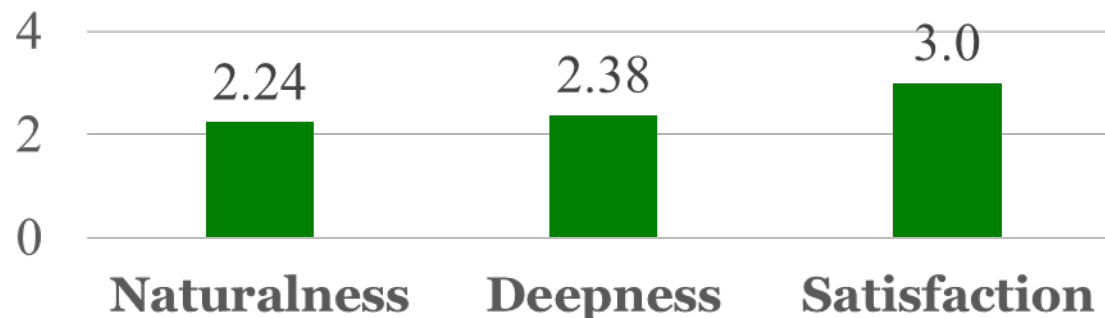
Topics	Support	Oppose
Casino bill	38	68
Capital punishment	158	183
Nuclear power plant reoperation	54	220

**Number of subjects :** 8 person

**Evaluation metric:** 5 point Likert scale (1-5)

- **Naturalness** - system response was natural or not
- **Deepness** - discussion was established or not
- **Satisfaction** - new knowledge was obtained or not

## Experimental result:





## Conclusion

- A debate system was proposed to help users to understand news deeply in the question answering system “NetTv”.
- User stance (support/oppose/neutral) and the reason (present /absent) were estimated using CNN and evaluated.
- Debate database was constructed by extracting arguments from Web data for three topics and debate was subjectively evaluated.

## Future works

- There are gaps between the system and user response due to the lack of value evaluation between them.
- The debate database must have many high quality arguments and they have to be converted from writing style to speaking style with unification of the characters as well as styles.

***Thank you!***